

The GEOID96 high-resolution geoid height model for the United States

D. A. Smith, D. G. Milbert

National Oceanic and Atmospheric Administration/National Geodetic Survey, 1315 East-West Highway, Silver Spring, MD 20910, USA
e-mail: dru@ngs.noaa.gov; Tel.: +1 301 713 3202; Fax: +1 301 713 4172

Received: 28 July 1997 / Accepted: 2 September 1998

Abstract. The 2 arc-minute \times 2 arc-minute geoid model (GEOID96) for the United States supports the conversion between North American Datum 1983 (NAD 83) ellipsoid heights and North American Vertical Datum 1988 (NAVD 88) Helmert heights. GEOID96 includes information from global positioning system (GPS) height measurements at optically leveled benchmarks. A separate geocentric gravimetric geoid, G96SSS, was first calculated, then datum transformations and least-squares collocation were used to convert from G96SSS to GEOID96.

Fits of 2951 GPS/level (ITRF94/NAVD 88) benchmarks to G96SSS show a 15.1-cm root mean square (RMS) around a tilted plane (0.06 ppm, 178° azimuth), with a mean value of -31.4 cm (15.6-cm RMS without plane). This mean represents a bias in NAVD 88 from global mean sea level, remaining nearly constant when computed from subsets of benchmarks. Fits of 2951 GPS/level (NAD 83/NAVD 88) benchmarks to GEOID96 show a 5.5-cm RMS (no tilts, zero average), due primarily to GPS error. The correlated error was 2.5 cm, decorrelating at 40 km, and is due to gravity, geoid and GPS errors. Differences between GEOID96 and GEOID93 range from -122 to $+374$ cm due primarily to the non-geocentricity of NAD 83.

Key words. Geoid · GPS · Datums · Reference systems · Gravity

1 Introduction

Geoid models over the land can be computed by three different techniques: gravimetric, astro-geodetic, and, most recently, by differencing ellipsoidal heights and orthometric heights on leveled benchmarks occupied by

global positioning systems (GPS). The gravimetric and the GPS benchmark techniques, in particular, are highly complementary. A gravimetric geoid model can provide very high resolution and superior local accuracies, but is subject to long-wavelength (>100 km) systematic effects due to error accumulation. A network of GPS heights on leveled benchmarks has higher accuracy over long distances, but cannot approach the spatial scale of gravimetric models. In theory, an optimal combination can be achieved through the integrated geodesy approach (e.g. Eeg and Krarup 1973; Milbert 1988). However, practical problems are seen when trying to apply integrated geodesy in some areas (Milbert and Dewhurst 1992).

Problems associated with data combination are highlighted in comparisons of GPS/level undulations with high-resolution geoid models in the US (Milbert 1991b, 1995). Systematic offsets, tilts, and localized discrepancies are apparent, with magnitudes ranging around ± 1.5 m. The source of a transcontinental tilt evident in earlier geoid models has been identified as the non-geocentricity of the NAD 83 reference frame, used to express the GPS ellipsoidal heights. However, even when one accounts for this tilt, a systematic bias between geoid undulation models and measured GPS/level undulations is seen at the 30–40-cm level. This bias is believed to be due to the offset of the NAVD 88 reference level with respect to the current best estimates of the global mean sea level (Burša 1995). With the NAD 83 tilt and the NAVD 88 bias removed, localized disagreements are still seen at the 15.6-cm RMS level. The disagreements are probably due to errors in the geoid model, as well as older, questionable data in statewide GPS surveys.

GEOID96 represents a new era of geoid models from the National Geodetic Survey of the United States (NGS). Previous models such as GEOID90 (Milbert 1991a) and GEOID93 (Milbert and Schultz 1993) were gravimetric, geocentric models of the geoid. Milbert (1995) explored the possibility of computing a “biased” geoid model which would directly relate GPS heights in

the NAD 83 frame to orthometric heights in the NAVD 88 vertical datum. That study led to the decision to produce both a gravimetric, geocentric geoid model (G96SSS) and a non-geocentric geoid model (GEOID96), as there were two distinct user groups, each with their own geoid needs. Pre-release results are found in Milbert and Smith (1996).

Section 2 discusses geoid computation from a global spherical harmonic model. EGM96 (Lemoine et al. 1997). Section 3 continues with the computation of G96SSS. In Sect. 4 the GPS/benchmark set is described, and the use of that benchmark set in the computation of GEOID96 is found in Sect. 5. In Sects. 6, 7, and 8 a comparison of GEOID96 and G96SSS to three other geoid models is made: the Canadian GSD95 model, the EGM96 model, and GEOID93, respectively. The results are discussed in Sect. 9 and the conclusions are summarized in Sect. 10.

2 The spherical harmonic model, EGM96

G96SSS was calculated using an established mathematical procedure, commonly referred to as “remove–compute–restore” (Schwarz et al. 1990). The procedure can be summarized as follows.

1. Remove a gravity anomaly field (spherical harmonic model evaluated at the geoid) from Helmert gravity anomalies based on local gravity measurements and digital elevation data. This leaves “residual gravity anomalies.” Faye anomalies are used to approximate Helmert gravity anomalies.
2. Compute “residual co-geoid undulations” by a spherical Fourier representation of the Stokes integral.
3. Restore a geoid undulation field (spherical harmonic model evaluated at the geoid) to the “residual co-geoid undulations” and add a primary indirect effect (computed from digital elevation data) to form the geoid undulations.

It can be seen that the spherical harmonic model of the Earth’s geopotential plays an important role in this computational procedure. A recently released geopotential model, EGM96 (Rapp and Nerem 1994; Lemoine et al. 1997) was chosen to be used in computing G96SSS.

EGM96 was preceded by five preliminary models (X01–X05), and all six models were evaluated by the special working group (SWG) of the International Geoid Service (IGeS). Reports of the working group were made available on the World Wide Web (Sideris 1996), and are available in Bulletin No. 6 of the IGeS (IGeS 1997). The model evaluations (Smith and Milbert 1997a, b) showed that EGM96 was a good model for the remove–compute–restore procedure in the US. While local discrepancies are pointed out, it was concluded that the long-wavelength structure of the model was the most important factor in the models, and that EGM96 was shown to have the best long-wavelength structure of any of the models, when compared with 2951 GPS/leveling

benchmarks in the US [in the ITRF94(1996.0) and NAVD 88 reference frames].

Traditionally, geoid undulations and gravity anomalies have been computed from geopotential coefficients by evaluating a series expansion at a point on the geoid (i.e. inside the crustal masses). However, Rapp (1997) shows that computation of an undulation from geopotential coefficients is erroneous, inside the crustal masses. The traditional geopotential evaluation, used in this paper, is valid if the geoid undulations and gravity anomalies are treated consistently. Detailed discussion is presented in the Appendix.

When computing geoid undulations one must also be cautious about the treatment of the permanent tides of the Earth. The three frequently used systems, mean, zero and non-tidal (tide free) all have advantages and disadvantages. The mean tide system retains masses external to the Earth, and is therefore incompatible with Stokes’ integral. For this reason, the mean tide system was not used. Although International Association of Geodesy (IAG) Resolution No. 16 (IAG 1984) recommends the use of the zero tide system, the preponderance of data available to the NGS is reduced to the non-tidal system. This includes the GPS ellipsoidal heights and NAVD 88 Helmert heights. Although arguments against the non-tidal system exist (e.g. Ekman 1996), consistency with data sources was preferred. As such, G96SSS and GEOID96 were produced in the non-tidal system.

3 The gravimetric geoid model, G96SSS

The method for computing the gravimetric geoid model, G96SSS, is based on the use of the one-dimensional (1-D) spherical Fast Fourier Transform (FFT) to evaluate Stokes’ integral. As described in Schwarz et al. (1990), a bandwidth-limited signal is needed for input to the FFT convolution. Such bandwidth limitation is closely approximated through applying terrain corrections (removal of high-frequency data) and removal of a spherical harmonic model (low frequency).

The FFT requires gridded data. Any gridding procedure is subject to aliasing in the presence of high-frequency signal. Therefore, one should remove as much predictable, high-frequency content as possible. For this reason, gridding is performed on complete (i.e. terrain-corrected) Bouguer anomalies (Milbert 1995):

$$\Delta g_{TB} = g + \frac{2\gamma_a}{a} \left[1 + f + m + \left(-3f + \frac{5}{2}m \right) \sin^2 \varphi \right] \times H - 3 \frac{\gamma_a}{a^2} H^2 - \gamma + A + C_P - 0.1119H \quad (1)$$

where

$$C_P = \frac{1}{2} G_\rho R^2 \iint_\sigma \frac{(H - H_P)^2}{l_0^3} d\sigma \quad (2)$$

$$A = 0.8658 - 9.727 \times 10^{-5}H + 3.482 \times 10^{-9}H^2 \quad (3)$$

$$l_0 = 2R \sin(\psi/2) \quad (4)$$

Δg_{TB}	terrain-corrected Bouguer gravity anomalies
g	surface gravity, tide-corrected, IGSN 71 system (mGal)
H	orthometric height, NAVD 88 datum (m)
γ	normal gravity on ellipsoid, GRS80 (Somigliana's formula) (mGal)
A	atmospheric correction (Wichiencharoen 1982) (mGal)
C_P	terrain correction (mGal)
φ	geodetic latitude, NAD 83 (86) datum
a	semi-major axis, GRS80 (6 378 137 m)
γ_a	normal gravity at equator, GRS80 (978 032.67715 mGal)
f	ellipsoid flattening, GRS80 (0.0033528106818)
m	$\omega^2 a^2 b / GM = 0.00344978600308$ (GRS80)
ρ	density of topographic masses ($2.67 \times 10^3 \text{ kg/m}^3$)
G	universal gravitational constant
R	mean radius of the Earth
ψ	spherical distance
l_0	horizontal distance approximation

The gridding algorithm uses a method of continuous curvature splines in tension (Smith and Wessel 1990) with tension parameter $T_B = 0.75$. This value was selected to minimize the impact of gravity errors in mountains on adjacent grid points without gravity data. The method is one which honors the data and does not display large oscillations in areas without data coverage. The product is a 2 arc-minute \times 2 arc-minute regular grid extending from 24 to 53°N and 230 to 294°E (66 to 130°W). Thus, the grid contains 871 rows and 1921 columns. All anomalies are pre-filtered by computing mean value and mean location of the anomalies in 2 arc-minute \times 2 arc-minute cells centered over the regular 2 arc-minute latitude and longitude intersections. This pre-filtering step is recommended by Smith and Wessel (1990) to reduce spatial aliasing effects prior to gridding.

It is necessary to convert the grid of terrain-corrected Bouguer anomalies into a grid of Helmert anomalies. As an approximation to Helmert anomalies, a grid of Faye anomalies is generated instead. Martinec et al. (1993) state that the approximation is associated with a linear relationship between free-air anomalies and elevation. An alternative approach to obtain a Helmert anomaly can be found in Martinec and Vaníček (1994), Vaníček and Martinec (1994), Martinec et al. (1996), and Vaníček et al. (1996). However, this approach entails large computational requirements. For this reason, the Faye anomaly was used in accordance with past NGS practice (Milbert 1991a, 1995; Milbert and Schultz 1993).

In light of the preceding discussion, the geoid undulation, N , is computed through the generalized Stokes' integral (Heiskanen and Moritz 1967, p. 102)

$$N = \frac{R}{4\pi\gamma} \iint_{\sigma} (\Delta g + C_P) S(\psi) d\sigma + \delta N_1 + N_0$$

$$= \bar{S}(\Delta g + C_P) + \delta N_1 + N_0 \quad (5)$$

where

$$S(\psi) = \frac{1}{\sin \frac{\psi}{2}} - 4 - 6 \sin \frac{\psi}{2} + 10 \sin^2 \frac{\psi}{2} - \left(3 - 6 \sin^2 \frac{\psi}{2} \right) \ln \left(\sin \frac{\psi}{2} + \sin^2 \frac{\psi}{2} \right) \quad (6)$$

and

$S(\psi)$	Stokes' function
$\bar{S}(\cdot)$	operator form of Stokes' integral
δN_1	indirect effect induced by using Helmert's second method of condensation $= \pi G \rho H^2 / \gamma$ (Wichiencharoen 1982)
N_0	degree zero geoid undulation induced by δGM and δW
δGM	GM (current best estimate) $- GM$ (GRS-80 ellipsoid value)
δW	W_0 (gravitational potential on geoid) $- U_0$ (normal potential on ellipsoid)

Therefore

$$N = \frac{R}{4\pi\gamma} \iint_{\sigma} (\Delta g_{TB} + 0.1119H) S(\psi) d\sigma + \delta N_1 + N_0 \quad (7)$$

In accordance with Eq. (7), the authors have restored the Bouguer plate term, $+0.1119H$, to the terrain-corrected Bouguer anomaly grid of Eq. (1), using a grid of 2 arc-minute \times 2 arc-minute mean elevations, yielding a grid of Faye anomalies.

Equation (7) is solved in a remove-compute-restore procedure using the "1-D" FFT formulation of Haagmans et al. (1993). First, the residual gravity anomalies are computed

$$\Delta g_+ = \Delta g_{TF} - \Delta g_0 - \Delta g_{2,360} \quad (8)$$

where

Δg_+	high-frequency part of gravity anomalies
Δg_{TF}	terrain-corrected free-air (Faye) gravity anomalies
Δg_0	degree-zero gravity anomaly induced by δGM and δW , at geoid
$\Delta g_{2,360}$	spherical harmonic model Δg , degrees 2 to 360, at geoid

Then the residual co-geoid undulations are computed row-wise (i.e. for each φ_p) by

$$N_{+, \varphi_p}(\lambda_p) = \frac{R \Delta \varphi \Delta \lambda}{4\pi\gamma} F^{-1} \times \left[\sum_{\varphi_q = \varphi_1}^{\varphi_n} F(S_{\varphi_q}(\Delta \lambda_{pq})) F(\Delta g_{+, \varphi_q}(\Delta \lambda) \cos \varphi_q) \right] \quad (9)$$

followed by the restoration of the spherical harmonic model undulations and the inclusion of the indirect effect

$$N = N_+ + N_{2,360} + N_0 + \delta N_1 \quad (10)$$

where

$\Delta\phi, \Delta\lambda$	grid spacing
F, F^{-1}	direct and inverse, one-dimensional, Fourier transforms
N_+	high-frequency part of geoid undulations
$N_{2,360}$	spherical harmonic model geoid undulations, degrees 2 to 360, at geoid
δN_1	indirect effect induced by using Helmert's second method of condensation
N_0	degree zero geoid undulation ($\delta GM, \delta W$)
n	number of rows in grid

and where the subscripts p and q indicate the fixed row and floating row respectively. Equation (9) must be evaluated for each row p , where $1 \leq p \leq n$.

The input grid, Δg_+ (870 rows, 1920 columns), had 50% zero padding on the east and west edges of the grid, but none on the north and south (870 rows, 3840 columns). This is because the 1-D FFT procedure only performs the FFT in the east/west direction, and thus padding is only needed on those two edges to eliminate the effect of cyclic convolution (Gleason 1990). No tapering of Δg_+ was performed, since the long-wavelength part ($\Delta g_{2,360}$) has already been removed. The final computation time for the 1-D FFT procedure over the conterminous US was 18 hours, 41 minutes (wall clock).

Figure 1 displays a shaded relief image of the G96SSS geoid height model. The geoid heights range from a low of -52.8 m in the Atlantic Ocean to a high of -7.7 m in the Rocky Mountains. As seen in earlier models (e.g. Milbert 1991a), significant short-wavelength structure is evident.

From the foregoing it is evident that the G96SSS gravimetric geoid model is derived from four data sources: point gravity, digital elevations, altimetrically derived gravity anomalies (Sandwell and Smith 1997), and geopotential coefficients. Over 1.8 million gravity points, altimetric, ship and terrestrial, went into the gridding. The data were a combination of NGS-held

data and quality-controlled data from the Defense Mapping Agency (now contained in the National Imagery and Mapping Agency, NIMA). Almost all of the gravity data were on non-monumented points, and due to the consequent uncertainty in elevation, the anomalies can be considered to have a nominal accuracy of about 1 mGal. The digital elevation data came primarily from the 30 arc-second point topography database, TOPO30, distributed by the National Geophysical Data Center (Row and Kozleski 1991). The 30 arc-second point data were originally derived from 1:250 000 scale maps, and are considered accurate to 30 m (50 m in the mountains). In addition, an elevation grid with a 1 km nominal spacing for the Canadian Rocky Mountains was made available by the Geodetic Survey Division, Geomatics Canada on a 30 arc-second \times 60 arc-second grid. This data was regridded to 30 arc-seconds \times 30 arc-seconds and inserted into the TOPO30 Digital Elevation Model (DEM) above latitude 51°N , and west of 250°E longitude. The 30 arc-second elevation set was used to compute both terrain correction (TCs), and 2 arc-minute \times 2 arc-minute mean elevations. And, as discussed earlier, the EGM96 model was used to compute the long-wavelength gravity anomaly and geoid height grids.

As mentioned in Sect 2, G96SSS was computed in a non-tidal system. To establish the exact level of G96SSS, a best-fitting global non-tidal ellipsoid was defined. This ellipsoid has no bearing on the final reference ellipsoid (GRS-80) used to reference the undulations, but is an altimetrically derived ellipsoid that best fits global mean sea level (within the geographic coverage of satellite altimeters). This concept is discussed thoroughly in Rapp (1995). At the time of our computations, an ellipsoid was used whose parameters are $a = 6\,378\,137.59$ and $1/f = 298.25722\dots$ (yielding a W_0 value of $62\,636\,855.726\text{ m}^2/\text{s}^2$, when combined with $GM = 3.986004418 \times 10^{14}\text{ m}^3/\text{s}^2$ and $\omega = 7\,292\,115 \times 10^{-11}\text{ rad/s}$). After releasing G96SSS, an improved set

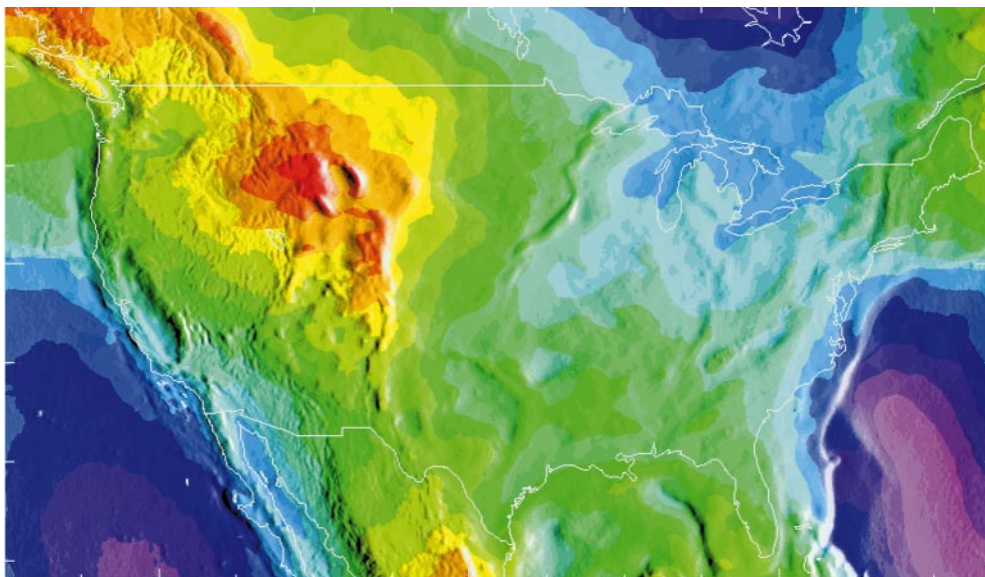


Fig. 1. Color image of G96SSS geoid height model

of best-fitting ellipsoid parameters was released from NIMA. In the computation of EGM96, the NIMA/GSFC team calculated the global best-fitting non-tidal ellipsoid parameters as $a = 6378136.46$ and $1/f = 298.25765$, and adopted the value of W_0 as $62\,636\,856.88 \text{ m}^2/\text{s}^2$. These differing values of W_0 imply that the geopotential on the level surface of G96SSS differs from that of EGM96 by $1.15 \text{ m}^2/\text{s}^2$. Spatially, the separation between these two surfaces may be computed by dividing by normal gravity (which yields a constant separation of 12 cm, to centimeter accuracy). This spatial separation of the two level surfaces requires that a correction be made to G96SSS. It is necessary to subtract 12 cm from the G96SSS values of obtain the geoid undulation between the best-fit global geopotential surface (the level surface which is best fitted by the EGM96 best-fit ellipsoid parameters) and the GRS-80 ellipsoid (both expressed in a non-tidal system). All further references to G96SSS, especially with regard to GPS/benchmark tests, will refer to this corrected surface.

The TCs were computed on a regular 30 arc-second grid by means of the FFT convolution of Sideris (1985), but altered by the FFT formulation for spherical distance of Strang Van Hees (1990), and exploiting Hermitian symmetry

$$C = KF^{-1}[F(H^2)R] - (2K)H \cdot F^{-1}[F(H)R] + KH^2R(0,0) \quad (11)$$

where

K	$G\rho/2$, with $\rho = 2.67 \times 10^3 \text{ kg/m}^3$
F, F^{-1}	direct and inverse, 2-D Fourier transforms
H	grid of elevations
H^2	grid of squared elevations
R	numerical transform of l_0^{-3} of Milbert [1995, Eq. (3.16)]

The grid of input heights (24–53°N, 230–296°E; 3480 rows, 7920 columns) was edge-tapered with a bell cosine function of 1° width, and then zero padded by an additional 3° width on all four edges. A 50% pad was felt wasteful, due to the rapid decay of the l_0^{-3} kernel. Tests with a 1-km-high, half-Bouguer plate confirmed this rapid decay. The final TC grid was extracted from the output of Eq. (11), spanning the region 25–51°N, 232–294°E. TCs at all gravity points in the region were then computed with bilinear interpolation.

The use of FFT for computing TCs requires geometric approximations, and may result in some lost information at high frequencies, when compared to a rigorous prism integration. Some tests were performed in areas of rugged topography to quantify the differences between TCs computed through FFT and through prism integration. In an area of southern British Columbia, TCs at 202 gravity measurement points were investigated in the area of 50–51°N and 235.5–237.5°E. Four different prism integration programs (both flat-top and slant-top types) were used as a control against the FFT terrain correction program. Using the 30 arc-second DEM, the FFT method was seen to agree with the prism

methods to $\pm 1 \text{ mGal}$ at points when the prism-integrated TC was below 30 mGal. In areas where the TCs exceeded 30 mGal (some 10 of all points in the test area), the FFT program had difficulty replicating the prism integration results, and was systematically lower than them by an average of 8 mGal. It must also be understood that both FFT and prism TCs were systematically smaller than the Canadian point TCs (provided with their gravity data) in southern British Columbia. Evaluation of residual gravity anomalies and geoid undulations in the northwest US showed significantly less signal when Canadian TCs were used, and thus Canadian point TCs were used north of 49°N in computing G96SSS.

Although the FFT had trouble replicating large TCs, computer resources and time constraints meant that the 30 arc-second DEM with the FFT procedure would be used for the bulk of GEOID96. Some independent tests in southern British Columbia have shown that a systematic 8-mGal error made on TCs greater than 30 mGal has a maximum localized effect of 8 cm. While significant, it is not large enough to ascribe the predominance of geoid errors to the FFT procedure. Of greater concern is the mismatch of both FFT and prism methods with the Canadian point TCs. It is probable that high-frequency information contained in the Canadian point values may not be contained in a 30 arc-second DEM. Limited prism integration tests with a 3 arc-second DEM in southern British Columbia indicate that TCs computed from a 3 arc-second DEM may nearly double the magnitude of TCs from a 30 arc-second DEM, and yet remain smaller than the Canadian provided TCs. Also, the questionable accuracy, unknown vertical datum, and sheer bulk of 3 arc-second DEM data that exists for the US and Canada made its use too cumbersome for GEOID96. Once all preliminary TC tests were complete, G96SSS was computed using 30 arc-second TCs, and represents geoid undulations, relative to GRS-80, centered at the origin of the ITRF94 (1996.0) reference frame.

This section closes with a discussion of the computation of GEOID96 in Alaska, Hawaii, and the Puerto Rico/Virgin Island area. Gravimetric geoids were produced for these three areas using terrestrial, ship and altimetric gravity data. All three of these areas have extensive oceanic boundaries, so altimetry was of great importance to all of their calculations. To provide edge control to each grid, altimetrically derived gravity anomalies (Sandwell and Smith 1997) were combined with ship anomalies in the ocean areas. However, detailed analysis of all ocean areas often gave precedence to the altimetry data, and (when necessary) some unreliable ship data was dropped, in preference for the altimetric data. Complex near-shore bathymetry can yield unreliable models of ocean tides, degrading the altimetrically derived gravity anomalies. Alaska, Hawaii, and Puerto Rico/Virgin Islands all have complex near-shore bathymetry and thus each area was thoroughly examined for how best to mix altimetry with ship data (in a way similar to that done with G96SSS). This involved weighting altimetry and ship data based on

sparseness of data, near-shore bathymetry, and altimetry/ship agreements. In all three areas a gravimetric, geocentric geoid was produced using the same techniques as G96SSS. However, because these three geoids were released as end-products, they were designated as GEOID96 models, rather than G96SSS. The relevant statistics for all three geoids are presented in Table 1, and are shown with the conterminous US values for G96SSS, for completeness.

4 The GPS benchmark data set

NGS is engaged in a project to establish a nationwide GPS network that is called the High-Accuracy Reference Network (HARN). HARN station spacing and accuracy is variable, ranging from 25- to 100-km spacing, and around 1 ppm accuracy. Many of the HARN points are also NAVD 88 benchmarks. Clearly, these form a powerful data set for accuracy assessment and improvement of geoid models. Figure 2 displays the locations of 2951 points that are optically leveled benchmarks with NAVD 88 Helmert orthometric heights, and which have GPS-measured ellipsoidal heights in the NAD 83 (86) reference system as of September 1996. The irregular distribution illustrates the

state-by-state approach to the surveying, and the different levels of state participation. Further detail on the HARNs can be found in Bodnar (1990), Milbert and Milbert (1994), and Milbert (1995). It must be emphasized that the HARN ellipsoidal heights are not as accurate as one would expect from continuously operating GPS receivers.

The positioning and navigation communities require coordinates in the NAD 83 (86) datum. While being primarily a horizontal, classical network, the NAD 83 (86) was controlled by VLBI and Doppler data sets, and can be considered 3-D. Snay (pers. commun. 1996), has computed the seven-parameter Helmert transformation from NAD 83 (86) to ITRF94(1996.0) with eight points common to both reference systems. The results are summarized in Table 2. The RMS of fit was 13 mm.

Figure 3 portrays the datum differences between NAD 83 (86) and ITRF94(1996.0) ellipsoidal heights referred to the GRS80 ellipsoid. It is seen that the non-geocentricity of the NAD 83 (86) reference frame induces a smooth, systematic difference in ellipsoidal heights. The values range from -0.28 to -1.64 m, and have an average tilt of about 0.3 ppm. Of particular note, this tilt is considerably smaller than the 1–2 ppm often seen in local intercomparisons. Summarizing, the ellipsoidal heights in the data set will contain vertical

Table 1. Statistics for GEOID96 in Alaska, Hawaii, and Puerto Rico/Virgin Islands

Name	Latitude boundaries	Longitude boundaries	Geoid grid spacing	Terrain correction grid spacing	Minimum geoid undulation [m]	Maximum geoid undulation [m]
Alaska	49–72°N	172–234°E	2' × 4'	30'' × 60''	-21.5	21.3
Hawaii	18–24°N	199–206°E	2' × 2'	6'' × 6''	-3.1	27.7
PR/VI	15–21°N	292–296°E	2' × 2'	6'' × 6''	-70.8	-34.4
G96SSS	24–53°N	230–294°E	2' × 2'	30'' × 30''	-52.8	-7.7

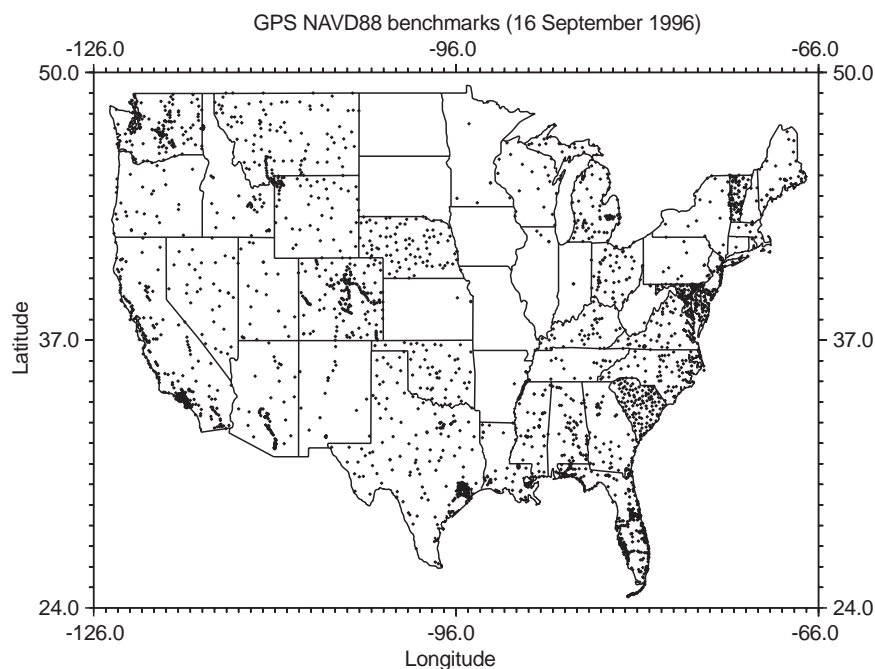


Fig. 2. 2951 leveled benchmarks with NAVD 88 Helmert orthometric heights and GPS ellipsoidal heights in the NAD 83 (86) reference system

Table 2. Transformation from NAD 83 (86) to ITRF94(1996.0)

ΔX	-0.9738	± 0.0261	m
ΔY	+1.9453	± 0.0215	m
ΔZ	+0.5486	± 0.0221	m
ω_x	-0.027 55	± 0.00087	arc sec
ω_y	-0.010 05	± 0.00081	arc sec
ω_z	-0.011 36	± 0.00066	arc sec
Scale	-0.007 78	± 0.00264	ppm

random error, and will have a long-wavelength systematic error component caused by NAD 83 (86) datum definition.

The NAVD 88 datum is expressed in Helmert orthometric heights (Heiskanen and Moritz 1967, p. 167), and was computed in 1991. The network contains over 1 million km of leveling at precision ranging from 0.7 to 3.0 mm/ $\sqrt{\text{km}}$ (Zilkoski et al. 1992). For relative geoid analysis in a local region, leveling can be considered nearly error free. Accuracy assessment of leveling at a national scale is more problematic. An interesting result is that shown in Figure 8 of Zilkoski et al. (1992). Two independent leveling data sets, that of Canada and that of the US, match at the 11-cm level or better at 14 points along the Canadian-US border. While repeatability is not a measure of accuracy, this suggests that the NAVD 88 vertical datum is not tilted.

The NAVD 88 datum was realized by a single datum point, Father Point/Rimouski, in Quebec. The strategy and the value of the constraint were based on a number of factors. However, the foremost requirement was to minimize recompilation of national mapping products. Thus, there is no guarantee that the NAVD 88 datum coincides with the normal potential, U_0 , defined by the GRS80 system, nor with the level of global mean sea level (Burša 1995). During the course of producing

G96SSS, a comparison was made between 2951 ITRF94(1996.0) GPS measurements at NAVD 88 leveled benchmarks and G96SSS geoid heights in the conterminous US. The results of that study indicate a mean bias for the NAVD 88 datum of -31.4 cm (± 15.6 cm). The sense of the sign is that the zero level surface of the NAVD 88 datum is below the best estimate of the Earth's global geoid (mean sea) level (which also does not correspond to $W = U_0$). This result is similar to that found by Balasubramania (1994). Also investigated was the magnitude of this bias for the eastern and western halves of the US, with results presented in Table 3. These results clearly indicate a nearly constant magnitude of the bias, further signifying that no tilt exists in NAVD 88.

5 Computing the conversion surface and GEOID96

Based on the forgoing results, it does seem clear that both the US national ellipsoidal height datum [NAD 83 (86)] and orthometric height datum (NAVD 88) are biased with respect to an ideal geocentric, best-fit, global reference system. In addition, the existence of 1–2 ppm, long-wavelength, systematic errors seen in this, and earlier, geoid/GPS benchmark comparisons suggests the need of an additional model component. Collocation of geoid/GPS benchmark misclosures will model the combined long-wavelength systematic errors in GPS leveling and the G96SSS model. Moreover, the combination of the error grid predicted by collocation with known datum relationships will allow development of a geoid model that will directly relate NAD 83 (86) to NAVD 88.

Modeling of the combined geoid, GPS and NAVD 88 errors begins by forming residuals, e , in the sense of

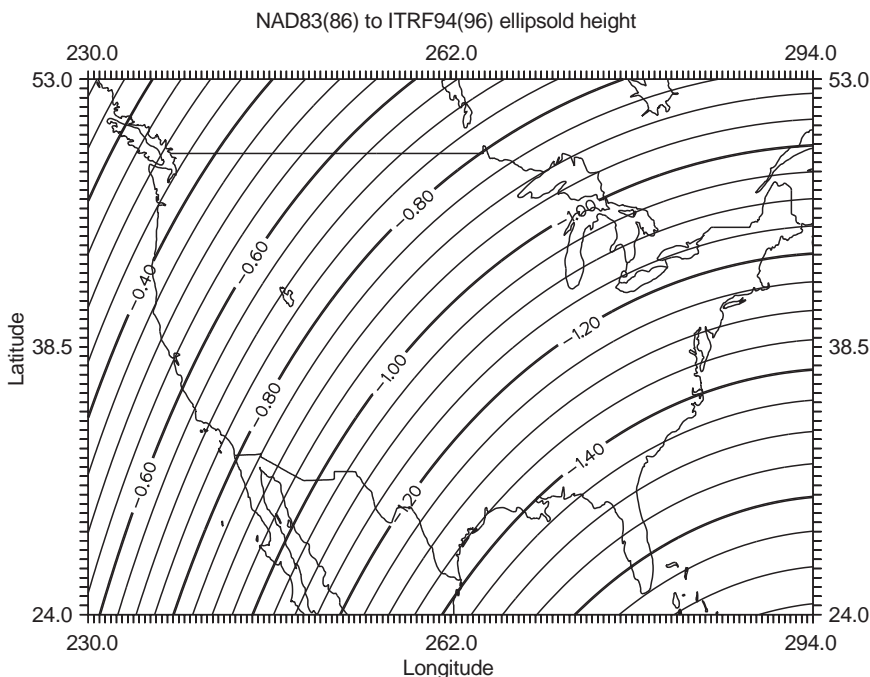


Fig. 3. Datum differences between NAD 83 (86) and ITRF94(1996.0) ellipsoidal heights referred to the GRS80 ellipsoid; contour interval = 4 cm

Table 3. NAVD 88 bias estimates for various areas in the US

Area	Number of GPS/BMs used	NAVD 88 bias estimate (cm)
Whole USA	2951	-31.4
East USA ($\lambda_E > 260^\circ$)	1431	-32.4
West USA ($\lambda_E < 260^\circ$)	1520	-30.5

e = G96SSS geoid height – (ITRF94(1996.0) ellipsoid height – NAVD 88 orthometric height).

Next, since collocation requires centered quantities (Moritz 1980, p. 76), a tilted-plane model was computed using the 2951 values of e , and this plane was removed from e to provide detrended residuals, which we denote as “ l ”. Results are shown in Table 4.

The mean offset of -31.4 cm should represent a good estimate of the bias in NAVD 88, since the GPS heights are transformed into ITRF94(1996.0) and do not therefore, contain the non-geocentric bias of NAD 83 (86). In addition, the lack of an east/west tilt indicates that there is no strong longitudinal, and, therefore, no strong height, dependence in the NAVD 88 datum bias (due to the strong east/west component of heights in the US).

Figure 4 displays an empirical covariance function fit to the covariance statistics of the detrended geoid residuals. The fit is made using a simple function of the form

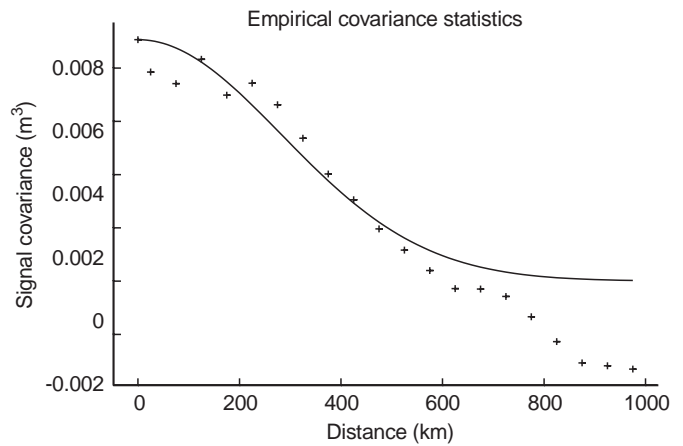
$$C = C_0 e^{(-d^1/L^1)} \quad (12)$$

where

- d the distance between points (km)
- L characteristic length (km)
- C_0 function variance (m^2)

In Fig. 4, the solid line indicates a function fit of $L = 400$ km and $C_0 = (0.095)^2 m^2$. The plus symbols display the covariance statistics. It is seen that the residual error is sizable and correlated over a long length scale. Thus, this error (or errors) is a smooth, slowly varying effect; but of a magnitude which exceeds our expectations of possible systematic effects in levelling and GPS. The dominant sources of this error are likely in the gravity data, in the digital elevations, and in the approximations to the true Helmert anomaly. Commission error in EGM96 is not a major contributor to the signal in Fig. 4, because commission errors over spatial scales of the grid size and smaller are removed by using surface gravity. See Smith and Milbert (1997b).

The detrended residual error, \hat{s} , is predicted on a 30 arc-minute \times 30 arc-minute grid using least-squares

**Fig. 4.** Empirical covariance function and detrended error statistics between ITRF94(1996.0)/NAVD 88 and G96SSS

collocation with noise (Moritz 1980, pp. 102–106). The prediction formula is

$$\hat{s} = C_{st}(C_{tt} + C_{nn})^{-1}l \quad (13)$$

where

- C_{tt} signal covariance between observations
- C_{st} signal covariance between predicted signal and observations
- C_{nn} covariance of random measuring errors, taken as diagonal and constant: $C_{nn} = s_0^2 I$
- l centered, detrended residuals

Before the grid is computed, the prediction process is iterated to establish a value of s_0^2 consistent with the residual misfit about the predictions of Eq. (13). It was found the RMS of residuals from the prediction step matched the assigned noise when $s_0^2 = (5.5)^2 cm^2$ was used for the $n = 2951$ points. This signal grid is seen in Fig. 5 and represents centered, detrended residuals between ITRF94(1996.0) GPS ellipsoid heights, NAVD 88 Helmert orthometric heights and G96SSS geoid heights. In viewing Fig. 5, it must be recalled that the GPS benchmarks used to develop this grid all lie within the US borders and that highs or lows in the oceans or in other countries are extrapolations, and are not reliable. To compute the total conversion surface between G96SSS and GEOID96, the trend reported in Table 4 is restored to the detrended signal grid, which is then added to the ITRF94(1996.0) to NAD 83 (86) grid (Fig. 3). This results in the final conversion grid (adjusted in sign to provide a subtractive conversion). Included in this conversion grid is the 12.0-cm correction that refers G96SSS to the global best-fit non-tidal system, as well as the 31.4-cm NAVD 88 datum offset. This combined grid is shown in Fig. 6.

Figure 6 portrays a conversion surface that, when subtracted from the G96SSS geoid model, produces the final geoid model, GEOID96. The GEOID96 model will directly relate NAD 83 (86) ellipsoid heights and NAVD 88 Helmert orthometric heights. Figure 6, like Fig. 5, is not reliable outside the boundaries of the

Table 4. Results from tilted plane fit

Mean offset	-31.4 cm
Tilt	0.06 ppm
Azimuth	178°
RMS about the mean	15.6 cm

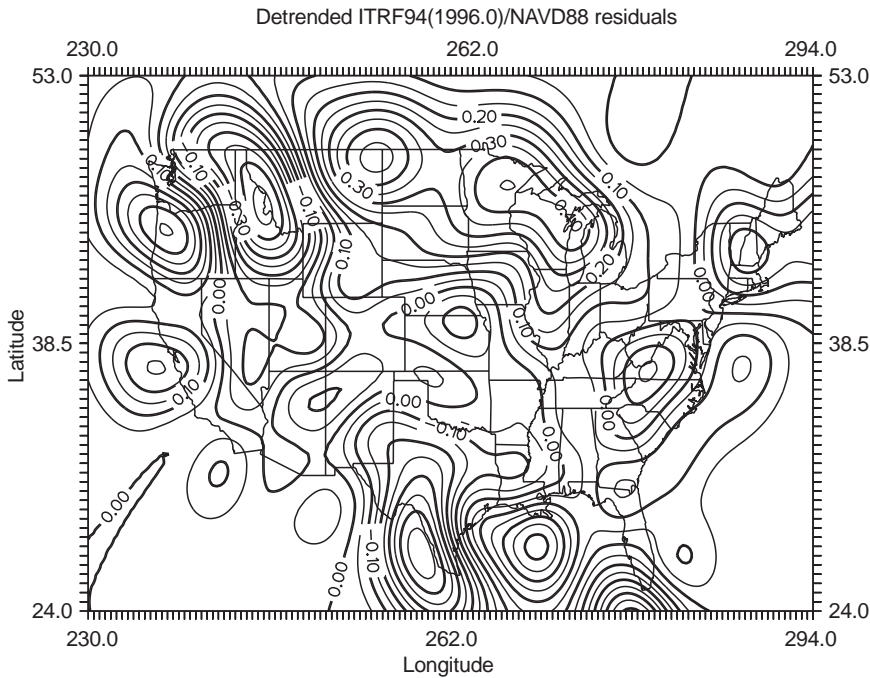


Fig. 5. Surface of ITRF94/NAVD 88 residuals with respect to G96SSS: contour interval = 5 cm

conterminous US. With this in mind, it is seen that the predominant effect is a tilt in the northwest-southeast direction [being mostly the transformation between ITRF94(1996.0) and NAD 83 (86)]. The conversion is smooth within the US (gradients under 4 ppm) and seldom exceeds 1 m.

To test the computational process leading to GEOID96, residuals were computed between the GEOID96 model and the measured geoid heights at the 2951 GPS benchmarks [this time, in NAD 83 (86)]. The RMS was 5.5 cm, with no offset or trends applied. Thus, the conversion process is seen to be successful.

Of particular interest, the covariance statistics for the GEOID96 model residuals were computed, and an empirical covariance function of the form of Eq. (12) was then developed. These results are portrayed in Fig. 7. Unlike Figure 4, which was plotted to a distance of 1000 km, Fig. 7 is only plotted to a distance of 100 km. At this closer scale, a drop is seen in the statistics from $C = (0.055)^2 \text{ m}^2$ at $d = 0$, down to $C = (0.025)^2 \text{ m}^2$ at $d = 5 \text{ km}$. This reduction is evidence of an uncorrelated (white-noise) process. For this reason, the empirical covariance function is fit to the remaining, correlated signal, yielding $L = 40 \text{ km}$ and $C_0 = (0.025)^2 \text{ m}^2$.

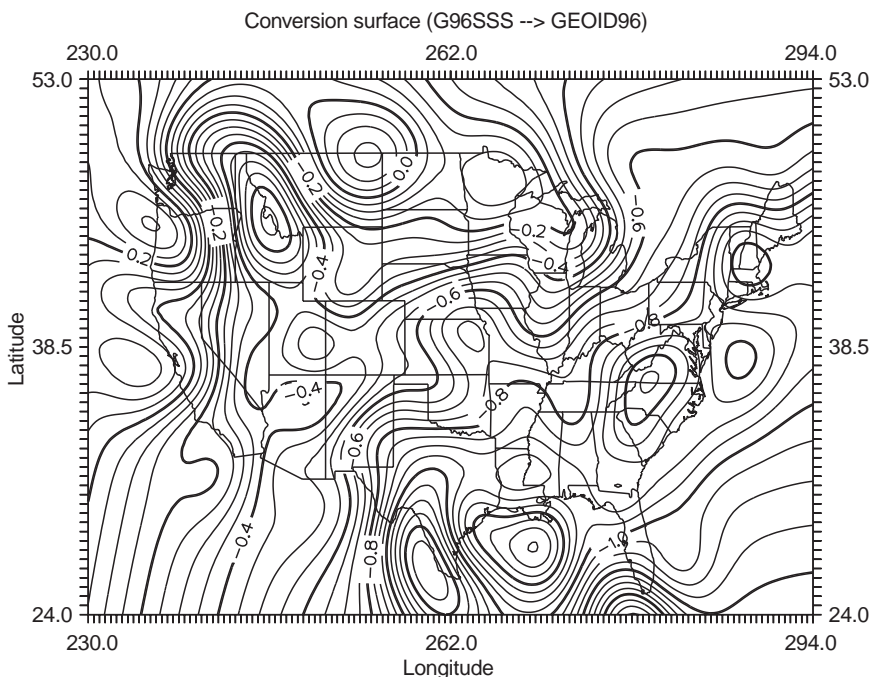


Fig. 6. Contour plot of conversion surface relating G96SSS to GEOID96; contour interval = 5 cm

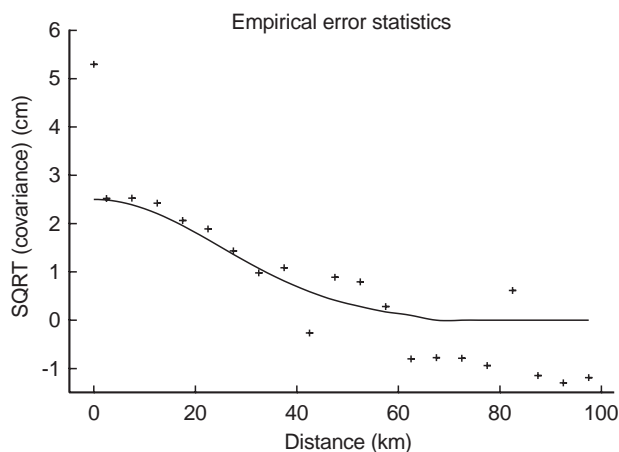


Fig. 7. Empirical covariance function and error statistics for GEOID96

A major source of the 5.5 cm of uncorrelated error is undoubtedly error in the GPS ellipsoidal heights. Both geoid height and leveling errors are correlated, and leveling is much too precise to contribute significantly to the value. Free-air gravity anomalies are well known in the US (1–2 mGal), and Parks and Milbert (1995) report a relationship of 3–4 mm of geoid error to a milligal of gravity error when using 3 arc-minute \times 3 arc-minute high resolution geoid models. In addition, the 5 cm uncorrelated error process is seen in both low- and high-elevation areas of the US. As mentioned earlier, error in the GPS ellipsoid heights of the data set was expected. The 5.5 cm magnitude is consistent with those field survey and reduction procedures.

The correlated error is difficult to interpret. It could be correlated GPS error, lower-order leveling error and geoid error from gravity data error and approximations. For example, standard errors of Second-Order leveling can range from 2.1 to 3.0 mm/ $\sqrt{\text{km}}$, or about 1.5 to 2.1 cm over 50 km (Zilkoski et al. 1992). It is notable that the GPS HARN network spacing is about 50 km. This may be related to the correlation length.

6 Comparison to Canadian geoid model, GSD95

While the goal of GEOID96 is to provide a high-resolution geoid model for the US, the computational

area extends about 400 km north into Canada to reduce the possibility of edge effects. This overlap provides a comparison area between US geoid models and Canadian geoid models. The most recent Canadian geoid, GSD95, was produced at the Geodetic Survey Division, Geomatics Canada (GSD/GC) in 1995 (Véronneau 1997). It is a model of geoid undulations in 5 arc-minute \times 5 arc-minute grid cells, extending south into the US to latitude 42°N, and was computed with slightly different procedures than GEOID96. To compare GEOID96 to GSD95, GEOID96 was interpolated onto the same grid as GSD95, and the overlapping areas of 42 to 53°N, 230 to 294°E were compared.

The first comparison is between G96SSS and GSD95, as they are both purely geocentric, gravimetric geoid models, with no GPS/benchmark information used in their computation. Figure 8 shows a plot of the differences G96SSS minus GSD95, with magnitudes ranging from -2.16 to $+0.21$ m (no bias removed). However, the north and south edges of this grid may be influenced by edge effects from the computation of G96SSS or GSD95. A general tilt exists in these differences of approximately 0.28 ppm at an azimuth of 67°. This translates into a roughly 1.25-m tilt in the east/west direction, over the span of 230 to 294° longitude. Even if one removes this trend, the large features (about 1 m in magnitude) covering Washington, Oregon, Idaho and western Montana remain. This large disagreement may be due to NGS and GSD/GC using different terrain correction sources between latitudes 42 and 49°N.

If one were to examine the differences between GEOID96 and GSD95, one would be influenced by the NAD 83/ITRF94 tilt, and the NAVD 88 bias. If, however, one were to remove these two known biases from GEOID96, that would leave a geoid model that is a combination of G96SSS and GPS/benchmarks in a geocentric reference frame. This model is denoted (for the purposes of this comparison) as “G96BM”. Figure 9 shows the differences G96BM minus GSD95. If one compares Fig. 9 to Fig. 8, it can be seen that the influence of GPS/benchmarks has changed the differences in a few areas. Specifically, notice that in Fig. 8 the agreement from central Montana (48°N, 255°E) to western Wisconsin (45°N, 270°E) is good, but the addition of GPS/benchmark data yields a worse agreement, as seen in Fig. 9. Leveling errors of this magnitude are not likely, however it is possible that there are errors in the GPS measurements in eastern Montana. In addition, the

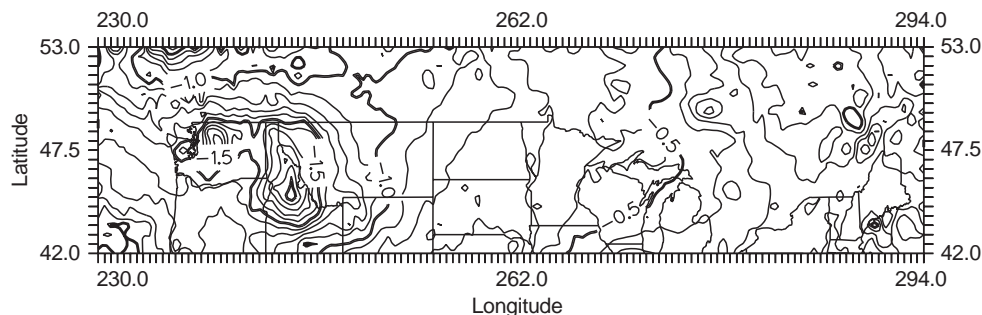


Fig. 8. Contour plot of differences between G96SSS and GSD95; contour interval = 10 cm

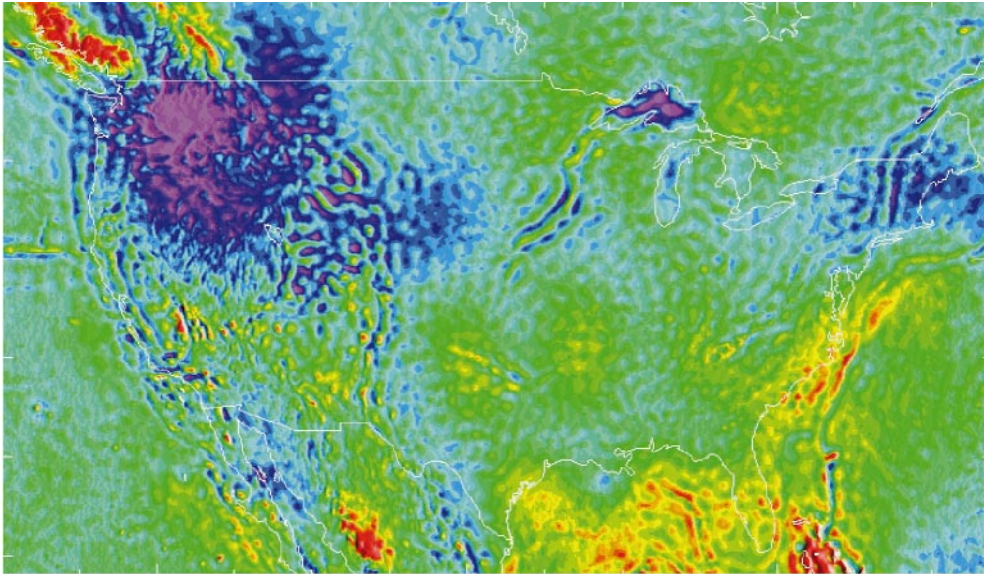


Fig. 10. Color image of differences between G96SSS and EGM96-implied geoid undulations (including height anomaly to geoid undulation correction)

GEOID96 to show the significance of the conversion surface. Figure 11 shows the differences GEOID96 minus GEOID93. These differences range in magnitude from -1.22 up to 3.74 m. The dominant feature of these differences is the northwest tilt due to the transformation between NAD 83 and ITRF94. The -31.4 -cm bias in NAVD 88 is another difference between GEOID93 and GEOID96. Significant changes in the terrain data in southern British Columbia, as well as gravity data in the Bahamas and the western Gulf of Mexico, are all responsible for the large changes seen in those areas. Finally, high-frequency differences in the entire ocean domain can be seen, which are due to the inclusion of altimetrically derived gravity anomalies in computing GEOID96.

Note the significant tilts induced across Washington State and Florida, due to corrections made in southern British Columbia and the Bahamas respectively. These tilts were seen when GEOID93 was compared to GPS on NAVD 88 benchmarks, and have now been significantly reduced in GEOID96 but, as seen in Fig. 6, these features are still far from eliminated.

9 Discussion

Figures 4 and 7 show the covariances of the differences between the GPS benchmarks and the G96SSS and GEOID96 model, respectively. The remarkable change between these two figures demonstrates the effectiveness

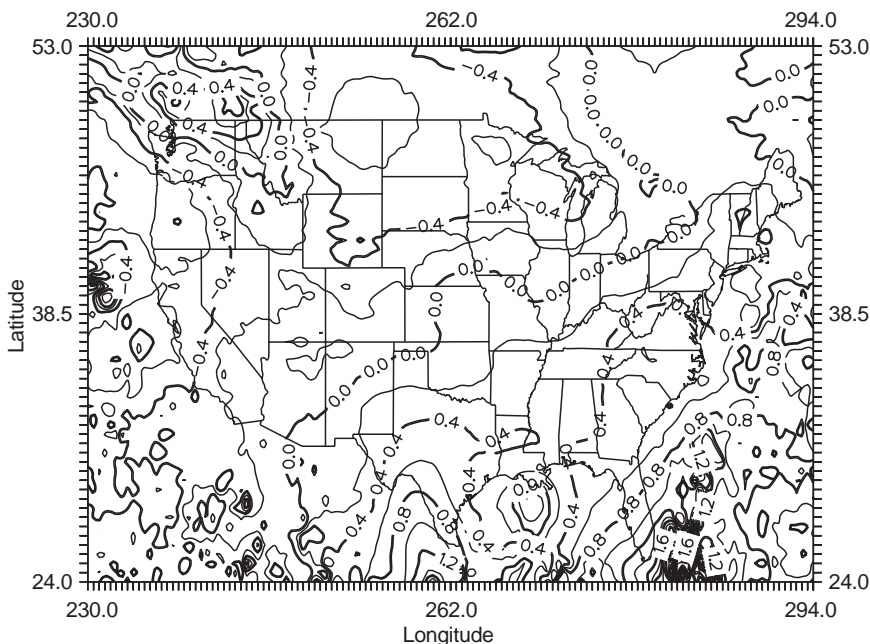


Fig. 11. Differences between GEOID96 and GEOID93; contour interval = 20 cm

of the collocation procedure. The detrended, correlated error was reduced from an RMS of 15.6 cm (Table 4), to an RMS of 2.5 cm, in the presence of 5.5 cm RMS Gaussian noise (Fig. 7). Further, this result was obtained from a very smooth empirical covariance function of $L = 400$ km. The error present in the GPS benchmark residuals, e (which have geoid model, NAVD 88, and GPS sources), fell into two distinct spectral domains: long-wavelength (~ 400 km) and short-wavelength (~ 40 km).

The long-wavelength error sources can be partially identified by inspection of the prediction grid. Since the conversion surface of Fig. 6 is dominated by the datum difference between ITRF94(1996.0) and NAD 83 (86), interpretation of the detrended, centered residuals of Fig. 5 is performed.

Figure 5 is only valid over land, with few GPS benchmarks contributing to its character in the central US. The dominant structure is an east–west tilt in the Pacific Northwest area (centered around 45°N , 245°E), which then slopes back down across western Montana. While the western part of Washington State is mountainous, the eastern part contains the Columbia Basin (47°N , 241°E) and is fairly flat, yet it shows greater tilt. In addition, other states which have equivalent or greater relief and which are well-sampled by GPS benchmarks, such as Colorado, show less structure. This discrepancy may be related to G96SSS model problems in southern British Columbia. However, with the addition of new digital elevation data, and the use of Canadian TCs in southern British Columbia, the source of geoid error in this area is elusive. The signal is much too large to be solely due to NAVD 88 leveling error.

The high in Fig. 5 centered over the northeast corner of Montana is likely due to GPS ellipsoid height error. The G96SSS model agrees well with the GSD95 model (Fig. 8) in the region. However, when the GPS benchmark data are included, then the high is generated (Fig. 9). Because of the very benign topography in the area, it is not realistic to assume that the agreement before the addition of GPS data was due to a common geoid error in both G96SSS and GSD95. More likely, G96SSS is reasonable in this area, and the collocation procedure has transferred long-wavelength systematic error from the GPS heights into the conversion surface and the GEOID96 model in that region.

On the other hand, note the tilt seen in the southern half of Florida in Fig. 5. This represents a reduction of the tilt seen throughout the entire state in Fig. 7 of Milbert (1995). During computation of the G96SSS model, it was found that the older, statewide tilt was caused by a lack of gravity coverage in the Bahamas region. The addition of satellite altimetry close to land in that region improved the geoid model in Florida. Recently, the remaining southern Florida tilt of Fig. 5 has been traced to ship gravity, incorrectly coded as land gravity data south of Florida. Therefore, this tilt in Florida exemplifies an improvement to the G96SSS model by the GPS benchmark data.

Next, consider the Gulf of Mexico. The sign of the tilt along southern Louisiana and the Gulf coast of Texas is

consistent with subsidence. While some areas along the Gulf coast are known to subside at over 1 cm/year, the overall magnitude in Fig. 5 is too large to be due solely to subsidence. The magnitude of the tilt along the west half of the Gulf of Mexico has been reduced from that seen in GEOID93 and Milbert (1995). This reduction was accomplished by the inclusion of altimetric gravity anomalies in the near-shore areas of the Gulf of Mexico, west of New Orleans. But, because of the introduction of altimetric anomalies, G96SSS disagrees with EGM96 in the Gulf of Mexico. However, the agreement with GPS/benchmarks was our primary goal. Again, the tilts in the Gulf coast region (Fig. 5) exemplify improvement to the G96SSS model by the GPS benchmark data.

A broad, 10–40-cm feature is seen in Fig. 5 near the Great Lakes, over Minnesota and Wisconsin. However, there are relatively few GPS benchmarks in this area. The GPS survey in Wisconsin dates from 1991, and was designed for horizontal accuracy requirements. While the mid-continent gravity high is in the region, it is a narrow, linear feature, and is perpendicular to the elongation of the broad feature. The density variation of the Great Lakes was not incorporated in the computations, but Martinec et al. (1995) show the effect to be in the vicinity of 1 cm. Nonetheless, it is suspected a portion of this feature is related to gravity data/reduction in Lake Superior.

In general, it can be seen that the majority of signal of Fig. 5 is commission error in G96SSS. It is highly likely that most of this commission error is due to data errors, as opposed to shortcomings in theory. One expects the greatest impact of approximations to be in the mountains. However, many features in Fig. 5 are found in areas of very low topographic relief. Certainly, a major feature is found in the mountainous northwest US, but the topography south of 40°N has equivalent roughness, even higher elevations (e.g. the Colorado Rockies), and does not show major features in Fig. 5. If use of a Faye anomaly causes large problems, then it should fail throughout the western US and not just in the northwest corner. This is not to say that the substitution of a Faye anomaly for a rigorous Helmert anomaly is ideal [Martinec et al. 1993, Eq. (34)]. Rather, the data errors are masking the effect of using Faye anomalies. At this time, no study exists which compares Faye anomalies to more elaborate Helmert anomalies at the geoid, or which compares GPS on benchmarks to a geoid computed from Helmert anomalies. The magnitude of a Faye anomaly approximation on the geoid is an open problem. It is expected that when the data errors are resolved one will see less commission error, and that the commission error will be more closely correlated to topographic relief.

In closing this section, a few remarks can be made on the computation technique for GEOID96. In one sense, the procedure may be viewed as a variant form of the integrated geodesy approach, where an empirical covariance function now substitutes for the computationally inaccessible cross-covariance matrix of the gravity data and the associated geoid. On the other hand, the empirical covariance function is quantifying systematic

error, rather than random error. The peak of the empirical covariance function is 9.5 cm (Fig. 4), whereas random error propagation of 1-mGal gravity data would probably result in 1–2 cm or better geoid height precision (Kearsley 1986; Parks and Milbert 1995). The collocation procedure used to compute GEOID96 was constructed to accommodate datum discrepancies and long-wavelength systematic errors. Large-scale systematic error in GPS height can be transferred into the GEOID96 model by this procedure. This is actually a desirable property, since users will establish subsequent coordinates using that same erroneous control. The systematic effects in GEOID96 will compensate, leading to accurate orthometric heights. This does imply that large-scale readjustments of the ellipsoidal height network will have to be accompanied by corresponding upgrades to GEOID96-type models.

10 Conclusions

A gravimetric, geocentric geoid model, G96SSS, was computed by means of a 1-D FFT convolution about the EGM96 spherical harmonic model of the geopotential. Comparison of 2951 ITRF94(1996.0) GPS benchmarks with NAVD 88 Helmert orthometric heights identified a 15.6-cm RMS variation about a tilted plane with a mean of -31.4 cm and a north/south trend of 0.06 ppm. If one considers ITRF94(1996.0) as geocentric, this indicates a vertical datum bias of -31.4 cm in NAVD 88. No east/west trend is evident in the NAVD 88 offset. In addition, essentially the same NAVD 88 offset is obtained by fits to regionally clustered subsets of GPS/level benchmarks. A simple, empirical covariance function with a long characteristic length, $L = 400$ km, was found to fit the detrended differences between the geoid model and the GPS benchmarks.

The least-squares collocation predictor led to the development of a smooth signal surface. Adding this surface to the -31.4 cm offset and 0.06 ppm trend, as well as the known ITRF94(1996.0) to NAD 83 (86) surface, yielded the final conversion surface. Removing this surface from the G96SSS geoid model produced a new geoid model, GEOID96. This model is relative to a non-geocentric ellipsoid, and relates the NAD 83 (86) datum to the NAVD 88 datum. The differences between GEOID96 and NAD 83 (86)/NAVD 88 GPS benchmarks have a 5.5-cm RMS. Analysis of the covariances indicates the 5.5-cm value is Gaussian noise, and is likely due to random error in the GPS ellipsoidal heights. The figure is more than double the error ascribed to geoid and/or leveling, and highlights the need to improve GPS ellipsoidal height accuracies. Evaluation of the covariances of the differences between GEOID96 and the GPS benchmarks indicates a geoid accuracy of 2.5 cm with an empirical covariance function characteristic length of $L = 40$ km.

Because of the smooth character of the collocation surface and the datum transformations comprising the conversion surface, the GEOID96 model retains the high relative accuracy of G96SSS, in wavelengths shorter than 400 km. The collocation procedure was

seen to remove systematic errors in the G96SSS model, such as poorly covered regions across the country and improperly coded gravity data in the Florida region. However, the technique was also seen to add error in locations with isolated, unreliable GPS heights, specifically in the area of Montana and North Dakota.

Acknowledgements. This study incorporates the contributions of numerous NGS employees involved in the creation and evaluation of the gravity, NAVD 88, and GPS data sets. The National Imagery and Mapping Agency (NIMA, formerly DMA) provided a major portion of the NGS land gravity data, and was instrumental in the creation of various 3 arc-second and 30 arc-second digital elevation data grids in use today. NIMA was also a partner in a joint project with Goddard Space Flight Center, that developed the EGM96 global geopotential model. The Geodetic Survey Division of Geomatics Canada supplied the GSD95 geoid height model, as well as gravity and digital terrain data for comparisons. Dr. Walter Smith, NOAA, provided altimeter-derived gravity anomalies used in the G96SSS and GEOID96 models. The authors wish to thank the reviewers for their constructive comments.

Appendix

The remove–compute–restore procedure requires the removal of spherical harmonic model (SHM) “gravity anomalies” prior to evaluating Stokes’ integral, and the restoration of SHM “geoid undulations” after evaluating Stokes’ integral. The computation of G96SSS used a well-known method of computing both the model gravity anomalies and geoid undulations at the location of the geoid; that is, in the masses of the Earth’s crust. However, computing geoid undulations from geopotential coefficients, anywhere below the surface of the Earth, is known to yield geoid undulations that are up to 3.5 m in error globally (Rapp 1997) and up to 1.4 m in error in the US. G96SSS was computed using an approximation of Helmert’s second method of condensation, and therefore one should consider deviation of SHM “geoid undulations” from true *co*-geoid undulations. Using the results of Rapp (1997), and removing the indirect effect, it is seen that in the US, SHM “geoid undulations”, computed at the location of the geoid, deviate from *co*-geoid undulations by up to 0.7 m. By extension, it may be assumed that SHM gravity anomalies computed in the same way also deviate from true Helmert gravity anomalies.

In the next section it is proven that the model gravity anomalies and model geoid undulations, evaluated within the masses of the Earth, are connected through Stokes’ integral. This will be called “analytical consistency”, and the level of approximations in this consistency is also discussed. In the final section the implications of using such model values are briefly discussed in the light of recent research (Rapp 1997).

Analytical consistency within the remove–compute–restore technique

Consider the following: compute a function, which will be called “model gravity anomalies”, at any location (r, θ, λ) internal or external to the Earth’s surface, from

disturbing potential coefficients, in a spherical approximation, by

$$\Delta g_{0,360}(r, \theta, \lambda) = \frac{GM}{r^2} \sum_{n=0}^{360} (n-1) \times \left(\frac{a}{r}\right)^n \sum_{m=-n}^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) + \frac{2}{r} \delta W \quad (\text{A1})$$

Similarly, “model geoid undulations” are computed as follows:

$$\begin{aligned} N_{0,360}(r, \theta, \lambda) &= \frac{GM}{r\gamma} \sum_{n=0}^{360} \left(\frac{a}{r}\right)^n \sum_{m=-n}^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) - \frac{\delta W}{\gamma} \\ &= N_0 + N_I + \frac{GM}{r\gamma} \sum_{n=2}^{360} \left(\frac{a}{r}\right)^n \\ &\quad \times \sum_{m=-n}^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) \\ &= N_0 + N_I + N_{2,360} \end{aligned} \quad (\text{A2})$$

where

GM	gravity–mass constant
a	equatorial scale factor
\bar{E}_{nm}	fully normalized coefficients of disturbing potential
r, θ, λ	spherical coordinates of evaluation point
$\Delta g_{i(j)}$	model gravity anomaly at degree(s) i (through j)
$N_{i(j)}$	model geoid undulation at degree(s) i (through j)
δW	$W_0 - U_0$
\bar{Y}_{nm}	$\bar{P}_{nm}(\cos \theta) \cos(m\lambda), \quad m \geq 0$
\bar{Y}_{nm}	$\bar{P}_{n m }(\cos \theta) \sin(m \lambda), \quad m < 0$
\bar{P}_{nm}	fully normalized Legendre functions

Now, consider the remove–compute–restore technique. Begin with the generalization of Stokes’ integral to an arbitrary reference ellipsoid (using Helmert’s second method of condensation). Assume that degree 1 terms of the geoid undulation are zero.

$$N = \frac{R}{4\pi\gamma} \int_{\sigma} \Delta g_H S(\psi) d\sigma + N_0 + \delta N_I \quad (\text{A3})$$

where

N	geoid undulation
Δg_H	Helmert gravity anomalies on the geoid
$S(\psi)$	Stokes’ function
N_0	degree zero geoid undulation
δN_I	indirect effect on geoid undulations due to condensation

The remove–compute–restore technique changes Eq. (A3) to

$$\begin{aligned} N &= \frac{R}{4\pi\gamma} \left[\iint_{\sigma-\sigma_v} (\Delta g_H - \Delta g_{0,360}) S(\psi) d\sigma \right. \\ &\quad + \iint_{\sigma_v} (\Delta g_H - \Delta g_{0,360}) S(\psi) d\sigma \\ &\quad \left. + \iint_{\sigma} (\Delta g_{0,360}) S(\psi) d\sigma \right] + N_0 + \delta N_I \end{aligned} \quad (\text{A4})$$

where $\Delta g_{0,360}$ represent gravity anomalies computed on the geoid from the SHM. In looking at Eq. (A4), the first term (a far-field effect) is assumed negligible; the second term represents the computational area of the FFT (σ_c); and the third term is replaced with the “restored” undulations, $N_{2,360}$, of the geopotential model. The fourth term is the degree zero term of the undulations, and the fifth is the indirect effect. It is the replacement of the third term that drives our concern for analytical consistency. Note that the degree zero terms of the undulation are accounted for in the N_0 term, and the degree 1 terms have been assumed to be zero, so that the $N_{2,360}$ term which replaces the third term above is derived from the degree 2 through the 360 geopotential coefficients.

Thus, to rigorously use the remove–compute–restore procedure, it will be proven that

$$N_{2,360} = \frac{R}{4\pi\gamma} \int_{\sigma} c(\Delta g_{0,360}) S(\psi) d\sigma \quad (\text{A5})$$

By proving this, the analytical consistency required between N and Δg , irrespective of whether they are evaluated within the masses or not, is shown.

Begin with the expression of potential in spherical harmonics. The Earth’s external gravity potential, W , may be expressed outside a sphere of radius “ a_1 ” by

$$\begin{aligned} W_{0,\infty}(r, \theta, \lambda) &= \left(\frac{GM_1}{r} \sum_{n=0}^{\infty} \left(\frac{a_1}{r}\right)^n \sum_{m=-n}^n \bar{C}_{nm} \bar{Y}_{nm}(\theta, \lambda) \right) \\ &\quad + \Phi(r, \theta) \end{aligned} \quad (\text{A6})$$

where $\Phi(r, \theta)$ is the contribution of centrifugal potential to the gravity potential.

Similarly, the normal gravity potential outside of a reference ellipsoid of equatorial radius “ a_2 ” may be written as

$$\begin{aligned} U_{0,\infty}(r, \theta, \lambda) &= \left(\frac{GM_2}{r} \sum_{n=0}^{\infty} \left(\frac{a_2}{r}\right)^n \sum_{m=-n}^n \bar{D}_{nm} \bar{Y}_{nm}(\theta, \lambda) \right) \\ &\quad + \Phi(r, \theta) \end{aligned} \quad (\text{A7})$$

Note that the centrifugal potentials are identical in Eq. (A6) and (A7) and therefore the disturbing potential (at a point outside both the sphere, $r = a_1$, and the ellipsoid, $a = a_2$) may be written as

$$T_{0,\infty}(r, \theta, \lambda) = \frac{GM}{r} \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^n \sum_{m=-n}^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) \quad (\text{A8})$$

where

$$\begin{aligned}\bar{E}_{nm} &= \bar{C}_{nm} - \left(\frac{GM_2}{GM_1}\right) \left(\frac{a_2}{a_1}\right)^n \bar{D}_{nm} \\ GM &= GM_1 \\ a &= a_1\end{aligned}\quad (\text{A9})$$

Now, define the gravity anomalies in the “spherical approximation” (Heiskanen and Moritz 1967, p. 89)

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r}T + \frac{2}{r}\delta W \quad (\text{A10})$$

and also define the geoid undulations (Heiskanen and Moritz 1967, p. 85)

$$N = \frac{T}{\gamma} - \frac{\delta W}{\gamma} \quad (\text{A11})$$

Applying these definitions to Eq. (A8) and truncating at degree 360 will yield Eqs. (A1) and (A2). Also, Stokes’ function can be written as a function of (non-normalized) Legendre polynomials, P_n (Heiskanen and Moritz 1967, p. 97)

$$S(\psi) = \sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos \psi) \quad (\text{A12})$$

In order to complete the proof of Eq. (A5), one must evaluate the right-hand side (*RHS*) of Eq. (A5) using Eqs. (A1) and (A12). The algebra is lengthy, but a few key steps are shown below. Begin with

$$\begin{aligned}RHS &= \frac{R}{4\pi\gamma} \iint_{\sigma} \Delta g_{0,360} S(\psi) d\sigma \\ &= \frac{R}{4\pi\gamma} \iint_{\sigma} \left(\sum_{n=0}^{360} \sum_{m=-n}^n \frac{GM}{r} \frac{(n-1)}{r} \left[\frac{a}{r}\right]^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) \right) \\ &\quad \times \left(\sum_{n'=2}^{\infty} \frac{2n'+1}{n'-1} P_{n'}(\cos \psi) \right) \cos \varphi d\varphi d\lambda \quad (\text{A13})\end{aligned}$$

First, make multiple use of the following rule of summations:

$$\left(\sum_i A_i \right) \cdot \left(\sum_j B_j \right) = \left(\sum_i \sum_j A_i \cdot B_j \right) \quad (\text{A14})$$

Then exchange summation and integration; change Eq. (A12) to refer to fully-normalized Legendre polynomials; expand the Legendre polynomials using the decomposition formula [Heiskanen and Moritz 1967, Eq. (1-82), p. 33]; make use of orthogonality relations (Heiskanen and Moritz 1967, Sect. 1-13); and finally, recognize that all terms where $n' > 360$ will drop out due to the orthogonality relations, so that

$$\begin{aligned}RHS &= \frac{R}{4\pi\gamma} \sum_{n=2}^{360} \sum_{n'=2}^{360} \sum_{m=-n}^n \sum_{m'=-n'}^{n'} \int_{\varphi=-\pi/2}^{\pi/2} \int_{\lambda=0}^{2\pi} \\ &\quad \times \left(\frac{1}{n'-1} \frac{GM}{r} \frac{(n-1)}{r} \left[\frac{a}{r}\right]^n \right) (\bar{E}_{nm} \bar{Y}_{nm}(\varphi, \lambda)) \\ &\quad \times (\bar{P}_{n'm'}(\sin \varphi) \cos m' \lambda \bar{P}_{n'm'}(\sin \varphi') \cos m' \lambda' \\ &\quad + \bar{P}_{n'm'}(\sin \varphi) \sin m' \lambda \bar{P}_{n'm'}(\sin \varphi') \sin m' \lambda') \\ &\quad \times \cos \varphi d\varphi d\lambda \quad (\text{A15})\end{aligned}$$

Then, separate the summation over m' into portions where $m' = m$ and $m' \neq m$; split the summation over n' similarly for $n' = n$ and $n' \neq n$; recognize that the primed latitude and longitude are constants; use the orthogonality relations found in Heiskanen and Moritz [1967, Eqs. (1-68) and (1-74), pp. 29–31]; assume r is a constant and move it outside the integrals; and simplify significantly to obtain

$$\begin{aligned}RHS &= \frac{R}{4\pi\gamma} \sum_{n=2}^{360} \sum_{m=-n}^n \left(\frac{1}{n-1} \frac{GM}{r} \frac{n-1}{r} \left[\frac{a}{r}\right]^n \right) \\ &\quad \times (4\pi \bar{E}_{nm} \bar{Y}_{nm}(\varphi, \lambda))\end{aligned}\quad (\text{A16})$$

Then that one arrives at

$$RHS = \frac{R}{r} \frac{GM}{r\gamma} \sum_{n=2}^{360} \left(\frac{a}{r}\right)^n \sum_{m=-n}^n \bar{E}_{nm} \bar{Y}_{nm}(\theta, \lambda) = \frac{R}{r} N_{2,360} \quad (\text{A17})$$

or

$$\frac{R}{4\pi\gamma} \iint_{\sigma} \Delta g_{0,360} S(\psi) d\sigma = \frac{R}{r} N_{2,360} \quad (\text{A18})$$

In this derivation, the assumption was made that r was a constant. Now it is seen that its value is the radial distance to the geoid undulation point. This adds an error which is dependent (to 1st order) on the location of the undulation evaluation point, and the flattening of the Earth. In addition, the value of R/r in Eq. (A18) adds a similar level of error. Therefore, it may be concluded that within errors $O(10^{-3})$, (A5) holds, proving the analytical consistency of SHM gravity anomalies and geoid undulations computed “inside the masses”.

An implication of this analytical consistency, when combined with the use of an FFT for geoid computations, is that commission errors of a spherical harmonic model (relative to true co-geoid undulations) at scales larger than the computational area are generally retained, while shorter-wavelength commission errors, within the area, are removed by the surface gravity data. This explains the behavior seen in Smith and Milbert (1997a).

Model values versus true values

As pointed out in Rapp (1997), Eq. (A2) is not the true representation of the geoid undulation in the masses, but is rather a sort of “height anomaly”. A modified version of Eq. (A2) was proposed in Rapp (1997), whereby the radial distance is effectively taken to the Earth’s surface, Eq. (A2) is used to compute a “height anomaly” at the surface, and a “Bouguer” correction is applied to compute the geoid undulation in the masses

$$N_{\text{geoid}} = \zeta_{\text{surface}} + \frac{\Delta g_B}{\gamma} H \quad (\text{A19})$$

where

Δg_B	the simple Bouguer anomaly at the surface of the Earth
$\bar{\gamma}$	mean normal gravity between the ellipsoid and telluroid
H	orthometric height
ζ_{surface}	height anomaly at the surface of the Earth
N_{geoid}	geoid undulation at the geoid

If one removes the indirect effect from Eq. (A19), then one arrives at co-geoid undulations (N^c)

$$N_{\text{geoid}}^c = N_{\text{geoid}} - \delta N_1 \quad (\text{A20})$$

The differences between Eqs. (A2) and (A20) are dominated by short-wavelength (<500 km) structures in the US. These structures, being smaller than the smallest dimension of the computational area, are removed by using surface gravity data itself. However, there is long-wavelength (0.02 ppm tilt) difference between Eqs. (A2) and (A20) as well. It may be assumed that a long-wavelength difference also exists in the gravity anomalies. It is not clear how these two long-wavelength errors interact.

Thus, two things may be concluded. First, the use of “in the masses” values for model gravity anomalies and model geoid undulations is mathematically sound within the context of Eq. (A4). Second, the use of a limited computational area ($\sigma_c \subset \sigma$) in Eq. (A4), combined with failure of a spherical harmonic geopotential model to yield the true values of Helmert gravity anomalies and true co-geoid undulations, will leave a small long-wavelength error in the gravimetric geoid computation. This small, long-wavelength difference is largely removed (up to the wavelength resolvable by the spatial scale of the geoid computation area) by the gravity measurements. Errors longer than this spatial scale will appear as a tilt across the region, correlated with elevation.

References

- Balasubramania N (1994) Definition and realization of a global vertical datum. Rep 427, Department of Geodetic Science and Surveying, The Ohio State University, Columbus
- Bodnar AN (1990) National geodetic reference system statewide upgrade policy. Tech Papers ACSM-ASPRS Fall Convention, 5–10 Nov pp A71–A82
- Burša M (1995) Report of Special Commission SC3, Fundamental constants. Int Assoc Geodesy, Paris
- Eeg J, Krarup T (1973) Integrated geodesy. Int Rep 7, The Danish Geodetic Institute, Copenhagen
- Ekman M (1996) The permanent problem of the permanent tide; what to do with it in geodetic reference systems? *Marees Terrestr* 125: 9508–9513
- Gleason DM (1990) Obtaining Earth surface and spatial deflections of the vertical from free-air gravity anomaly and elevation data without density assumptions. *J Geophys Res* 95(B5): 6779–6786
- Haagmans R, de Min E, van Gelderen M (1993) Fast evaluation of convolution integrals on the sphere using 1-D FFT, and a comparison with existing methods for Stokes’ integral. *Manuscr Geod* 18(5): 227–241
- Heiskanen WA, Moritz H (1967) *Physical geodesy*. Freeman WH, San Francisco
- IAG (1984) *The geodesist’s handbook 1984*. *Bull Geod* 58(3)
- International Geoid Service (1997) *The Earth gravity model EGM96: testing procedures at IGes*. IGes bull 6, Politecnico di Milano, Italy
- Kearsley AHW (1986) Data requirements for determining precise relative geoid heights from gravimetry. *J Geophys Res* 91(B9): 9193–9201.
- Lemoine FG, Smith DE, Kunz L, Smith R, Pavlis EC, Pavlis NK, Klosko SM, Chinn DS, Torrence MH, Williamson RG, Cox CM, Rachlin KE, Wang YM, Kenyon SC, Salman R, Trimmer R, Rapp RH, Nerem RS (1997) The development of the NASA GSFC and NIMA joint geopotential model. Gravity, Geoid and Marine Geodesy, Int Symp 117, Tokyo, 30 Sept–5 Oct 1996 (GraGeoMar 1996). Springer, Berlin Heidelberg New York. <http://cddis.gsfc.nasa.gov/926/egm96/egm96.html>; <http://www.nima.mil/geospatial/products/GandG/wgs-84/geos.html>
- Martinec Z, Vaníček P (1994) Direct topographical effect of Helmert’s condensation for a spherical approximation of the geoid. *Manuscr Geod* 19(5): 257–268
- Martinec Z, Matyska C, Grafarend EW, Vaníček P (1993) On Helmert’s 2nd condensation method. *Manuscr Geod* 16(5): 417–421
- Martinec Z, Vaníček P, Mainville A, Véronneau M (1995) The effect of lake water on geoidal height. *Manuscr Geod* 20(3): 193–203
- Martinec Z, Vaníček P, Mainville A, Véronneau M (1996) Evaluation of topographical effects in precise geoid computation from densely sampled heights. *J Geod* 70(11): 746–754
- Milbert DG (1988) Treatment of geodetic leveling in the integrated geodesy approach. Rep 396, Department of Geodetic Science and Surveying, The Ohio State University, Columbus
- Milbert DG (1991a) GEOID90: a high-resolution geoid for the United States. *EOS* 72(49): 545–554
- Milbert DG (1991b) Computing GPS-derived orthometric heights with the GEOID90 geoid height model. Tech Papers of the 1991 ACSM-ASPRS Fall Convention, Atlanta, 28 Oct–1 Nov. American Congress on Surveying and Mapping, Washington, DC, pp A46–A55
- Milbert DG (1995) Improvement of a high resolution geoid height model in the United States by GPS height on NAVD 88 benchmarks. New geoids in the world, IGes bull 4, International Geoid Service, Milan pp 13–36
- Milbert DG, Dewhurst WT (1992) The Yellowstone–Hebgen Lake geoid obtained through the integrated geodesy approach. *J Geophys Res* 97(B1): 545–558
- Milbert DG, Schultz D (1993) GEOID (the National Geodetic Survey geoid computation program). Geodetic Services Division, National Geodetic Survey, NOAA, Silver Spring, MD
- Milbert DG, Smith DA, (1996) Converting GPS height into NAVD 88 elevation with the GEOID96 geoid height model. Proc GIS/LIS ’96 Ann Conf, Denver, 19–21 Nov. American Congress on Surveying and Mapping, Washington, DC, pp 681–692
- Milbert KO, Milbert DG (1994) State readjustments at the National Geodetic Survey. *Surv Land Inf Syst* 54(4): 219–230
- Moritz H (1980) *Advanced physical geodesy*. Herbert Wichmann, Karlsruhe
- Parks W, Milbert DG (1995). A geoid height model for San Diego County, California, to test the effect of densifying gravity measurements on accuracy of GPS-derived orthometric heights. *Surv Land Inf Syst* 55(1): 21–38
- Rapp RH (1992) Computation and accuracy of global geoid undulation models. Proc Sixth Int Geod Symp Satellite Positioning, Columbus 17–20 Mar. The Ohio State University, Columbus, pp 865–872
- Rapp RH (1995) Equatorial radius estimates from TOPEX altimeter data. In: *Festschrift Erwin Groten*, Institute of Geodesy and Navigation, Univ. FAF Munich, D-85577 Neubiberg
- Rapp RH (1997) Use of potential coefficient models for geoid undulation determinations using a spherical harmonic repre-

- sensation of the height anomaly/geoid undulation difference. *J Geod* 71(5): 282–289
- Rapp RH, Nerem RS (1994) A joint GSFC/DMA project for improving the model of the Earth's gravitational field. Presented at Joint Symp Int Gravity Commiss/Int Geoid Commiss, Graz, Sept
- Rapp RH, Pavlis NK (1990) The development and analysis of geopotential coefficient models to spherical harmonic degree 360. *J Geophys Res* 95: 21 885–21 911
- Rapp RH, Wang YM, Pavlis NK (1991) The Ohio State 1991 geopotential and sea surface topography harmonic coefficient models. Rep 410, Department of Geodetic Science and Surveying, The Ohio State University, Columbus
- Row LW III, Kozleski MW (1991) A microcomputer 30-second point topography database for the conterminous United States, User Manual, Vers 1.0. Natl Geophys Data Center, Boulder, Co
- Sandwell DT, Smith WHF (1997) Marine gravity anomaly from GEOSAT and ERS-1 satellite altimetry. *J Geophys Res* 102(B5): 10 039–10 054
- Schwarz KP, Sideris MG, Forsberg R (1990) The use of FFT techniques in physical geodesy. *Geophys J Int* 100: 485–514
- Sideris MG (1985) A fast Fourier transform method for computing terrain correction. *Manuscr Geod* 10(1): 66–73
- Sideris MG (1996) Reports to the special working group of the International Geoid Service. <http://www.ucalgary.ca/~sideris/SWG/SWG.html>
- Smith DA, Milbert DG (1997a) Evaluation of preliminary models of the geopotential in the United States. IGeS bull 6, International Geoid Service, Milan. http://www.ngs.noaa.gov/PUBS_LIB/betatest.html
- Smith DA, Milbert DG (1997b) Evaluation of the EGM96 model of the geopotential in the United States. IGeS bull 6, International Geoid Service, Milan. http://www.ngs.noaa.gov/PUBS_LIB/egm96.html
- Smith WHF, Wessel P (1990) Gridding with continuous curvature splines in tension. *Geophysics* 55(3): 293–305
- Strang Van Hees G (1990) Stokes formula using fast Fourier techniques. *Manuscr Geod* 15(4): 235–239
- Vaniček P, Martinec Z (1994) The Stokes–Helmert scheme for the evaluation of a precise geoid. *Manuscr Geod* 19(2): 119–128
- Vaniček P, Sun W, Ong P, Martinec Z, Najafi M, Vajda Pl, ter Horst B (1996) Downward continuation of Helmert's gravity. *J Geod* 71(1): 21–34
- Véronneau M (1997) The GSD95 geoid model for Canada. Proc Int Symp Gravity, geoid, and marine geodesy (GRAGEOMAR 1996), Tokyo 30 Sept–5 Oct 1996. Springer, Berlin Heidelberg New York
- Wang YM, Rapp RH (1993) Gravimetric geoid undulation computations in the Gulf of Mexico. Special contractual rep to Jet Propulsion Laboratory, California Institute of Technology, Pasadena
- Wichiencharoen C (1982) Fortran programs for computing geoid undulations from potential coefficients and gravity anomalies. Int rep, Department of Geodetic Science and Surveying, The Ohio State University, Columbus
- Zilkoski DB, Richards JH, Young GM (1992) Results of the general adjustment of the North American Vertical Datum of 1988. *Surv Land Inf Syst* 52(3): 133–149